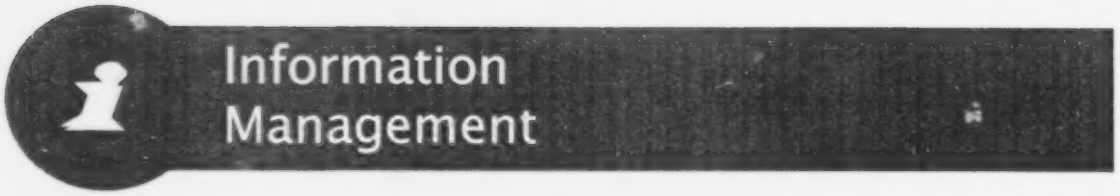

Naming Conventions for Electronic Documents

August 2005

Alberta
Government

Produced by

Information Management, Branch
Government and Program Support Services
Alberta Government Services
3rd Floor, Commerce Place
10155 – 102 Street
Edmonton, Alberta, Canada
T5J 4L4

Office Phone: (780) 422-2657
Fax: (780) 427-1120

Web site:
www.im.gov.ab.ca

©Government of Alberta

ISBN 0-7785-3685-8

CONTENTS

Contents i

1. Introduction 1

2. Electronic Workspaces 2

The Technology Environment 2
Directory and Folder Structure 3

3. Elements of Document Names 5

EIM or Controlled Repository Allows Shorter Names 6
Use of Properties Allows Shorter Names 6
Properties Metadata 7
EIM Environment Metadata 7
EIM Metadata Concerns 7
Properties Metadata Concerns 7

4. Using Naming Convention Elements 8

5. Decisions Required 11

Order 11
Separators 11
Abbreviations 12
Use of Headers and Footers 12
Versions 12

Appendix 1: Examples 13

Example #1: 13
Example #2: 13
Example #3: 13
Example #4: 14

1. INTRODUCTION

Many ministries have identified the need for consistent naming practices for electronic content items specifically documents in directories. Establishing naming conventions that are used consistently across an organization or an organizational unit will improve access to information in documents stored in an electronic work environment and in electronic information management (EIM) applications. Although there are many methods to electronically search for documents, users still tend to browse through directory structures looking for the documents they require.

Implementing good document names can support the following objectives:

- Facilitate better access to and retrieval of electronic documents.
- Allow sorting of documents in logical sequence (e.g. version #, date).
- Help users to identify the items they are looking for easily and support the ability to recognize the content of a document from a list.
- Help keep track of versions of documents.

The naming practices included in this guide apply to electronic documents created in a desktop environment and stored in shared or common directory structures. They are meant to assist ministries and business units to develop coherent and consistent approaches for naming their electronic documents.

These naming conventions should be part of the organization's defined business rules for managing shared electronic workspaces and should be used in conjunction with:

- Managing Shared Electronic Workspaces (Non–EIM Environment) Business Rules (being drafted)
- Electronic Information Management (EIM) Business Rules and Practices

105-07508/

2.
ELECTRONIC WORKSPACES

There are several factors to consider when determining appropriate naming practices. The most effective solutions will depend on the context in which documents are stored.

The Technology Environment

The structure of directories, folders and documents is nominally the same whether you are working on a standalone PC, a standalone server, a linked server or a controlled repository such as an electronic information management (EIM) application.

However, there are subtle differences in these workspaces that may influence the way in which directories, folders or documents are named. Each type of workspace has its own capabilities for naming and for additional indexing. They also handle information about the objects metadata in a slightly different manner.

There are several typical environments, including:

Environment	Storage	Taxonomy and Metadata
Standalone PC	Documents filed with applications or in partitioned drive.	Personalized identification scheme. Usually no taxonomy – document properties not used.
Standalone Business Unit Server	Documents filed in section-based directories for business documents. *Personal directories may be separate and include works in progress, undeclared records as well as personal documents.*	Each sub-unit may have its own taxonomy or folder structure. Personal folders tend to be many levels and hard to navigate. Document properties may be used to further index and support retrieval.
Shared Server-Decentralized Control	Documents filed in a shared directory environment as well as Business Unit directories. *Personal directories may be separate and include works in progress, undeclared records as well as personal documents.*	Multiple taxonomies – shared directory may be expanded by any business unit. Document properties may be used to further index and support retrieval.

Environment	Storage	Taxonomy and Metadata
Shared Server-Centralized Control	Documents filed in a shared directory environment as well as business unit directories. *Personal directories may be separate and include works in progress, undeclared records as well as personal documents.*	Common taxonomy in a shared environment; guideline or template provided for business unit structures. Naming conventions in place for documents and/or certain properties fields must be completed.
Shared Server with EIM Repository	Documents filed in a controlled repository for all declared business records. Shared workspace may be within the repository or remain undeclared *Personal directories provided for works in progress, undeclared records and personal documents.*	Controlled taxonomy. Metadata clearly defined and part of the classification/declaration of documents

Note that the better the capacity of the electronic environment to handle information about documents, e.g. document metadata itself, the less information will have to be captured in the naming structure.

Directory and Folder Structure

Directory and folder names by their nature will contain information that defines the document content. For example, a folder marked Accounting – Accounts Payable means that the documents contained in the folder need not contain the words accounts or accounting. Simply the name of the account or vendor may be sufficient.

Accordingly, when there are multiple levels of folders, the names of each of these folders need to be taken into account when deciding which file naming elements are required and whether or not to establish abbreviations for naming elements.

Several points to remember:

Location. Relying on the folder structure is only useful as long as the document continues to reside in the same folder. Should a document need to be moved or reclassified it may need to be renamed.

Name Length. In standard desktop applications, the entire path and file name are treated as the name of an item. If the complete name is longer than 255 characters, the system will truncate it, and important sections of the name may be lost.

Web Documents. If a document is to be posted on a web site, the version for posting may need a different name than the draft or other versions. (If a document is to be linked to a web site, the name of the document must be shorter. In RedDot, the government standard web content development application, there is a limit of 100 characters, including subdirectories. Internet file naming conventions may require a different approach, depending on the development application. For Internet files, it is useful to use shorter names and rely on metatags and metadata that support search engine queries.)

3.
ELEMENTS OF DOCUMENT NAMES

Document names should contain enough information to properly describe the contents of the document. However, keeping titles short will help users to quickly identify and retrieve accurate information. The following suggestions aim to strike a balance between creating usable, meaningful titles and brevity.

When determining the elements to be used and the order of the elements, it is critical to determine how the users will be retrieving the documents and design the naming convention appropriately. It may also be necessary to differ the structure and order of name elements from one directory to the other. The naming structure of documents in different folders needs to be based on user retrieval requirements.

For example, a "Meetings" file, may need to have the date of the meeting as the first element, thus ensuring all material related to one meeting date are sorted together. For example:

- X:\Research Services\ Staff Meetings\ 2005_01_04_Agenda.doc
- X:\Research Services\ Staff Meetings\ 2005_01_04_Background.doc
- X:\Research Services\ Staff Meetings\ 2005_01_04_Minutes & Decisions.doc
- X:\Research Services\ Staff Meetings\ 2005_02_05_Agenda.doc
- X:\Research Services\ Staff Meetings\ 2005_02_05_Minutes & Decisions.doc

However, in a "Publication" file, the date would only be required after the name of the publication and the version number. For example:

- X:\Research Services\Publications\Training Calendar_V01_2005_01_12
- X:\Research Services\Publications\Training Calendar_V02_2005_01_12
- X:\Research Services\Publications\Training Calendar_V03_2005_01_14

No matter which technology environment or folder structure is currently in place, the naming structure should include some or all of the following seven elements.

Name Elements	Example	Description of Need or Use
Title	Literature Reviews	The first three elements (title, version and date) are usually needed to facilitate searching for the document and the display of like documents in a logical order.
Version Number	V01	
Date (publication date, version date, or logical date relevant to the document)	2005_05_31	
Author or Creator	skessler	The middle elements (author, business unit, type) may be needed, depending on the business requirements. They can be helpful in identifying the controller of the document.
Business Unit/Program	Research	
Type (e.g. report, memo, letter)	RPT	
File Extension	.doc	The last element (file extension) is provided by the application in which the document has been created, and is always last. It is important to remember that this element should not be altered.

EIM or Controlled Repository Allows Shorter Names

If a shared electronic workspace environment has the capability to capture and search metadata, then the names of documents can be considerably shorter. For example, if the document author and business unit can be captured automatically by an EIM, a naming convention might contain only the following elements:

- Title
- Version Number
- Date

Use of Properties Allows Shorter Names

Other document elements can be captured as metadata elements. In MS Word, metadata elements can be captured in the "Properties" of the document (found under "File" in the main tool bar).

Properties Metadata

Relying on metadata in "Properties" can make information easier to find. Additional subject information and keywords can be added so that quick searches (not full text) will identify the documents.

EIM Environment Metadata

Within an EIM system, an interface to your application will be relied on heavily to acquire information to describe the document without user intervention. These metadata elements include such items as date created, date revised, version number, author, organization, type of document, subjects and others. The retrieval process will rely on searching these metadata fields.

In this environment, the actual name of the document can be relatively short, including only the title of the document. However, keep in mind that others who have not created the document may be browsing for information that is useful to them. As a result, including some additional elements in the title could be helpful (even though they may be captured in metadata) to facilitate browsing and retrieval.

EIM Metadata Concerns

In an EIM environment, it is important to consider how the system-generated attributes of the document (object title or number) are handled. In some cases, the EIM system may substitute the object title for the document title given to the document when it is saved as a new version. In these cases, the actual name of the document may include both the system-generated object title and the given title in combination.

Properties Metadata Concerns

If a previously created document is used to create a new document, the "Properties" of the new document retain the properties of the original document. In particular, the author will not be changed from the original. In these cases, "Properties" should be updated to reflect the new document.

It is useful to know that you can encourage users to complete the document properties by turning on the prompt for document properties within the application e.g. MS Word – Tools Options – Save.

This prompt only occurs the first time an individual saves a document.

4.
USING NAMING CONVENTION ELEMENTS

Some suggested practices around the use of each element are identified in Table 1.

Table 1
Use of Naming Convention Elements

Element	Text	Explanation
Title	Free Form	Think of the user/recipient and how they might search for and retrieve the document. Use a title that clearly describes the content of the document. The name should be intuitive to those who need to access it but not excessively long. Use standard abbreviations that have been agreed on by the organization and/or business unit. When creating multiple versions, use the same title for all versions. Consider if the sorting order is significant e.g. title first? date first? For very large documents where the component parts are stored separately, the title may also contain the "Part" of the document (e.g. Title_Part#_version, etc.). Part may be the chapter or section of the document.
Version	V01	Use a two-digit version number to facilitate sorting and display of documents in a logical order. If many versions or minor changes need to be tracked, an alternate format for version number could be V01-1, with the next substantive version being V02 followed by minor version changes as V02-1 and V02-2.
Date	2005-02-28	This is a significant date relevant to the document (e.g. version or published date). It is not the system-generated date that the software automatically updates each time a document is saved. Use the metric (international) date standard YYYY_MM_DD to facilitate the sorting and display of files in a logical order (see Alberta Data Standard). Use of the date should be defined by business needs, e.g. what is the most useful date for retrieval purposes. It may well be simply the year, e.g. for recurring reports.

Element	Text	Explanation
Creator/Author	sjones	Use first initial and last name. While this information may be correctly captured in the desktop document profile, it can be used to avoid confusion where multiple people are creating different versions of the document.
Business Unit/Program	IMB	Depending on the use of the information, this could be the business unit or program area. Be cautious of using branch names as these are subject to continual change. Program areas (which are business functions) change less often (e.g. they may reside in different branches or divisions or ministries over time). If branch names are used, they may be meaningless as soon as the next reorganization occurs. Identifying the program area can be useful in identifying the controller of the information. It may also be useful where there are committees with multiple sub-committees dealing with commonly titled documents.
Type		Document type can help reduce the length of the title. Where this information is captured as a metadata element, it is not likely to be needed in the document name. Do not include document type if its location (folder) identifies the type. The following type codes are examples. A full list needs to be developed that is consistent with GoA metadata standards currently under development. In most business units only a subset of these standard abbreviations may be needed. It is best to keep to a list of ten or less so that people can easily remember the codes.
	AGD	Agenda
	AGR	Agreement
	ARS	Action Request
	BRN	Briefing Note
	CPA	Cover Page
	CON	Contract
	DFT	Discussion Draft
	EXA	Example or Sample
	FRM	Form

Element	Text	Explanation
Type (cont'd)	GRA	Grant
	IDX	Index
	LTR	Letter
	LST	List
	MAC	Macro
	MEM	Memo
	MIN	Minutes
	MTG	Meeting
	NTS	Notes
	PLN	Plan
	PAP	Paper (e.g. research paper, discussion paper)
	POL	Policy
	PRS	Presentation
	PRC	Procedure
	RPT	Report
	SCH	Schedule
	SPE	Speech
	SUM	Summary
	SUP	Supplement
Extension	.doc .ppt .xls .pdf (etc.)	An extension must be provided. It is not optional. Use the appropriate extension for the application in which the document has been created. Extensions are automatically generated and attached to the end of the document name by the application. The extension will be used by the system to determine which application was used to create the document. Do not add different extensions or text after the extension.

5.
DECISIONS REQUIRED

Each organization will have to set consistent standards on:

- what naming elements are to be used and their order
- the use of a consistent separator
- the use of abbreviations
- what constitutes a new version number
- what to include in the header/footer of the document

Order

While not all components will necessarily be needed, the components should be placed in a consistent order within each folder. The actual order will depend on the structure of the folders that contain the document as well as the retrieval requirements of the users.

There are general rules of sorting that need to be considered. These rules include:

- "Nothing" comes before something. For example, "policy" will come before "policy disc", "agd" will come before "agenda."
- Numbers will be sorted from the left so "zero filling" may be required.

Sorting without zero-filling	Sorting with zero-filling
100	0100
1000	0110
1001	0120
110	1000
120	1001

Separators

To facilitate sorting, a consistent separator should separate all elements of the naming structure. The most common separator is the use of the underscore (_). However, other means of separating elements could be used, including a dash (-) or space. Spaces are not practical as a separator since there will be spaces within the elements themselves (e.g. document name). Remember that file names cannot include any of the following characters: forward slash (/), backslash (\), greater than sign (>), less than sign (<), asterisk (*), question mark (?), quotation mark ("), pipe symbol (|), colon (:) or semicolon (;).

It is important that a standard separator be adopted.

Some web-based applications interpret underscores differently.
In these cases, substituting dashes for underscores in the document name could be considered.

Abbreviations

Consistently used abbreviations can keep the name length manageable. However, if abbreviations are not standardized, they can be very difficult to interpret after the passage of time. A list of acceptable abbreviations should be developed for each organization.

If a code for a business unit is used, a common list of codes should also be developed to ensure consistent naming within the organization.

Use of Headers and Footers

Some of the components discussed in this guide could also be included in the header or footer of a document. This will facilitate easy location identification from printed documents.

Versions

There is no one recognized approach to identify a version. In many cases, when substantive changes have been made to the document, a new version would be created. However, this will be dependent on the information content and business context.

In the case where multiple team members are making changes to the same version of the document, the team members will have to determine when to create a new version.

APPENDIX 1: EXAMPLES

Example #1:

Using the full naming convention, this document could be named as follows:

Document Naming Conventions_V09_2005-08-31_mgamache_IMB_PRC.doc

In this form, the name includes all seven elements:

TITLE_VERSION_DATE_AUTHOR_ORGANIZATION_TYPE_EXTENTION

Example #2:

In the following example, the full convention is used, except that the document type (AGD for agenda) is included in the title and is therefore not a separate item in the document name.

2005-02-03_IMAC AGD_V01_skessler_IMB.doc

Example #3:

In some cases, the folder structure can carry some of the burden of the document name. For example, a business area has a project folder to which only certain people have access privileges. Inside that folder there is another folder for meetings and inside that is a folder for minutes. All drafts and accepted versions of minutes of each meeting are kept here. The folder structure would be "Project Name\Meetings\Minutes."

In this case, the naming convention may consist of:

- Meeting Date_Version_Author.doc
- 2005-02-15_V01_jjones.doc
- 2005-02-15_V02_ssmith.doc
- 2005-02-15_V02_llima.doc
- 2005-02-15_Approved_llima.doc

In this case, the folder structure (e.g. Project Name\Meetings\Minutes) carries a significant amount of information. The title of the document is the meeting date. However, should the document be moved to another folder, the context would be lost.

Example #4:

In an environment with a capability to search on metadata elements the name of this document could be simplified to:

Document Naming Conventions_V09_2005-08-31.doc

The reason for this simplification is the other elements would be captured in the document metadata. The version number and date are still included in the name to promote easier access to those who may not know (or remember these elements) when conducting a search for the document.